International Education Journal, 2006, 7(2), 189-206.
ISSN 1443-1475 © 2006 *Shannon Research Press*.
http://iej.com.au

Percentage population plots: A proposition for a new strategy for data analysis in comparative education

Pawel Piotr Skuza
School of Education, Flinders University *pawel.skuza@flinders.edu.au*

One of the issues facing educational research workers today is the determination of the similarities and differences between countries and cultures in the factors that influence educational outcomes. The author of this article proposes a new approach to this problem. Usually when countries are compared, the complete student samples are taken into consideration. At the same time, there are differences between countries with regard to their educational policies towards high or low achieving students as well as the effects of different student characteristics on the educational outcomes for those groups. Population Percentage Plots propose a new way of comparing the effects across the whole range of performance of groups of students.

Cross-national research, secondary data analysis, science achievement, comparative education, high and low achieving students

INTRODUCTION

Since the mid-1960s there have been substantial developments in the provision of secondary and higher education not only in the developed countries of the world, but also in many developing countries. The marked expansion of secondary education and the growth of universities have placed heavy financial burdens both on the wealthy nations and also on those nations that have growing financial commitments for infrastructure development to cater for a rapidly expanding population. The demand for accountability from the education sector has led to the introduction of different international testing programs that have undertaken surveys to assess student achievement at different levels of education. The international testing programs conducted by the International Association for the Evaluation of Educational Achievement (IEA) and those conducted by the Organisation for Economic Cooperation and Development (OECD) have provided valuable information for comparisons of the average levels of achievement between countries.

Initially these testing programs were committed to undertaking multivariate analyses to identify the factors that influenced educational achievement both across countries and within countries. The between country comparisons were undertaken in a search for the factors that had strong effects on educational outcomes. These analyses were limited by a lack of appropriate statistical procedures that enabled the teasing out of the factors, which operated at different levels of analysis, namely the student, classroom, school, region and country. However, while these problems have gradually been resolved, an under emphasis has emerged on the accurate estimation of the mean level of achievement of a national education system, without concern for the spread of scores in educational achievement and attitudes and the modelling of factors that influence the variation in scores both within and between countries. Some countries have sought to undertake multilevel and multivariate analyses of the data collected at a particular level of education within a country, and to publish the results of such analyses separately in national reports. However, there has been a noticeable absence of analyses that have examined change over time, across educational levels, and between kindred countries. As a consequence there has been little development of an understanding of the factors operating to influence student learning both within and across countries. Moreover, there has been little if any analyses conducted to examine how these factors have changed as a consequence of the marked expansion that has occurred in education. Of particular concern is that, there would seem to be a lack of interest in the performance of the very able students on whom the future of each nation must depend, particularly in the fields of science, mathematics and information and communications technology. In addition, there has been a lack of recognition of the significant role of such attitudes as perseverance, and interest in mathematics and science that required the accurate estimation of attitudinal data at the individual and sub-group levels.

It is the purpose of this article to develop a strategy for the examination of high and low performing sub-groups of students, both with respect to their achievement and their attitudes towards education, so that a greater understanding of the factors that influence both achievement and attitudes can be advanced.

Furthermore, this article is written at a time when each country is not only concerned with issues associated with "education for all" and "equality of educational opportunity", but is also very dependent on the development of talent, to support and advance the economic, scientific and technological development of the country.

CONCERN FOR THE DEVELOPMENT OF ABILITY

Theories of human learning indicate that students of the same age can be in different stages of cognitive development and can have different cognitive abilities. However, there seems to be a gap between the theories that highlight the variability among students and reports from achievement surveys that focus mostly on a whole national population and report national mean values and estimates of population, rather then sub-group, effects.

Furthermore, there is a second interesting issue. Different countries have different policies with regard to the allocation of more or less resources to help the higher achieving students. Two countries for example, that seem to differ in this matter, are Iran and the Republic of Korea. Iran took part in the Third International Mathematics and Science Study in 1999 and was classified in the 31st position out of 39 countries in Science achievement with an average scale score of 448 (3.8)[1] significantly below the international average of 488 (0.7) (Martin, 2000, p.32). Similarly, in Physics, Iran was classified in the 33rd position out of 39 countries with an average scale score of 445 (5.7) significantly below the international average of 488 (0.9) (Martin, 2000, p.99). However, students from Iran, who took part in the International Physics Olympiads (IPhO) were at the top of the international competitors, and Iran's best student got first, eleventh, third, seventeenth, second and third position in IPhO in 1997, 1998, 1999, 2000, 2001, and 2002 respectively (IPhO websites). In contrast, the Republic of Korea was high in TIMSS 1999 Science and Physics achievement (fifth with a score of 549 (2.6), and fourth with a score of 544 (5.1)) and the best students from Korea also did well in the International Physics Olympiads and got 56th, third, ninth, third, 42nd and ninth position in successive IPhOs. It is also worthy of mention that the total number of participants each year in IPhO was between 265 and 350.

In summary, according to the TIMSS 1999 study, on average the general population of students from Iran did not perform well when compared with the top achieving countries. On the contrary,

[1] Standard errors appear in parentheses.

in the International Physics Olympiads participants from Iran were in the top level of rankings. However, in both cases the Korean students were high in the cross-national rankings. From this comparison of Korea and Iran, it might be concluded that both countries strongly supported their more able students, but for different reasons Iran's students on average were not as high as Korea's.

These two reported findings indicate that further research into the different levels of students' achievement may provide very interesting information. The main purpose of this article is to present some ideas, which may form the beginning of a new strategy of data analysis that allows comparison of the impact of particular factors on student achievement and attitudes across countries and across different performance subgroups.

There are several important questions that guided the development of the strategy discussed and that may influence the direction of this approach to analyses in the future. Consequently, this initial introduction to the proposed method is based on these questions. As a short introduction to the method it can be said that it applies the simple principle of ordered subgroups selected according to the level of performance to examine the change in the estimated metric regression coefficients for successive subgroups, and to detect a pattern of change in the metric regression coefficients as an indicator of the change in relationship across different performance subgroups.

Although some interesting patterns and conclusions are presented in this article, the proposed method clearly needs further development.

At this stage it should also be noted that all analyses were done using data from the first Science Survey within the Programme for International Student Assessment (PISA) conducted in 2000. However, Science was not the highly tested subject on the PISA 2000 surveys, that was on this occasion Reading, with a lesser emphasis on Mathematics and Science. Only on PISA 2006 was Science the highly tested subject, while Mathematics occupied this position in PISA 2003. In addition, it is of interest to note that programmable macros in SPSS were used extensively in the data analyses. The PISA dataset provides two kinds of estimates of science scores: a weighted likelihood estimates (WLE) and a set of plausible values that resulted from a conditioning process. Little is known about the effects of the WLE procedure on the spread of scores and the estimation of the achievement levels of high performing students. In all analyses presented in this article only WLE estimates are used, therefore it has to be pointed out that any findings from this article relate to these estimates of achievement outcomes for samples collected in the PISA survey for each country under investigation.

Because of the novelty of the proposed method and the necessity for further development, collection of SPSS and Excel files, which were used, can be readily available for verification and request by e-mail (pawel.skuza@flinders.edu.au).

ELABORATION OF THE PROBLEM AND AN INTRODUCTION TO THE METHOD.

Question 1: Do the same relationships hold across different student performance levels as apply across the total student performance group for each country's sample?

A graph for different subgroups of students from Australia's PISA 2000 sample on the horizontal axis is presented in Figure 1. The line goes from a group of the top five per cent achievers to the top ten per cent and so on through the 100 per cent and bottom 90 per cent and to the bottom five per cent.

The vertical axis shows an unstandardised or metric regression coefficient (b), which was calculated between the so-called 'Warm estimate score' or 'weighted likelihood estimate' (WLE) for Science achievement and Sex of student for each achieving subgroup. The variable Sex of student was coded '1' for female, '2' for male. The standard deviations of scores for successive subgroups change markedly across sub-groups, and as a consequence, correlations and standardized regression coefficients cannot be compared across subgroups, although metric regression coefficients can be meaningfully compared.

Obviously, it is possible that an unstandardised regression coefficient between Science achievement regressed on Sex of student can be close to zero and sex does not significantly influence science achievement when considered for the whole student sample. However, there is a significant positive relationship when the sample is gradually restricted to the higher achieving students and a significant negative relationship for lower achieving students. This fact, about boys doing better than girls in the higher achieving groups and worse when considering the lower achievers, is not unexpected.

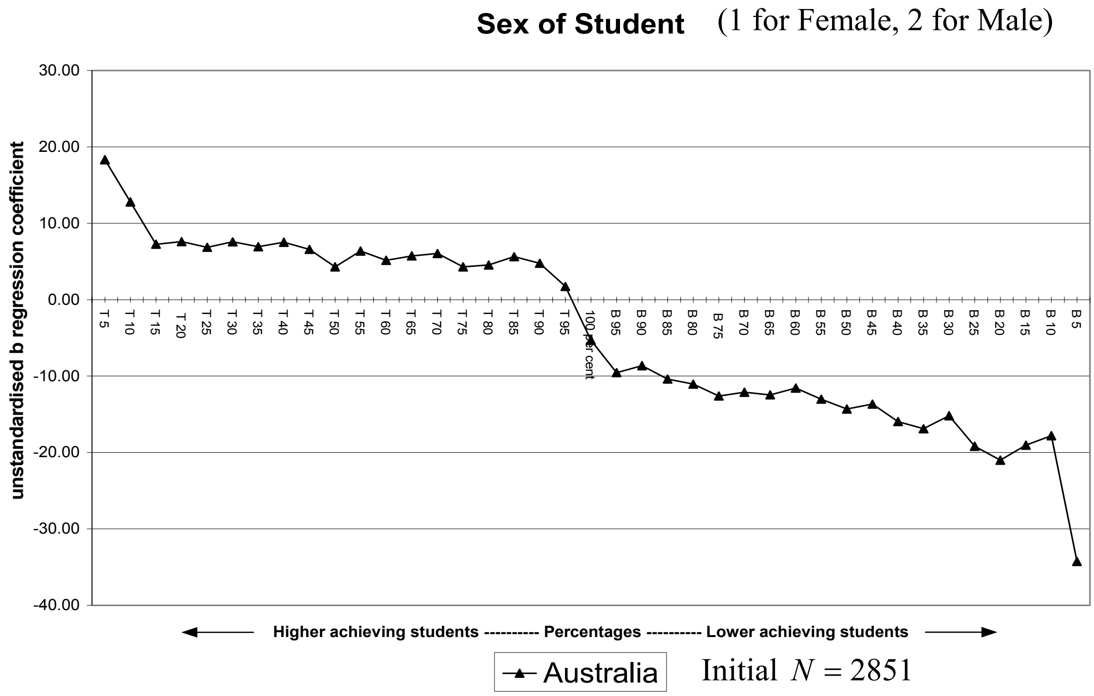


Figure 1 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable Sex of student for different Science achievement subgroups from Australia (Source file A1.1).

Figure 2 and Table 1 are presented below to assist with an explanation of the analysis carried out and the graph drawn in Figure 1. In Table 1 the results from Excel are presented with all b regression coefficients, standard errors and significance tests for Australia for the regression of Science achievement on the variable Sex of student for all subgroups examined. Although, more graphs like that in Figure 1 are presented in this paper, all additional information like that in Table 1 is not included but is available online in the appropriate Excel files through an AutoFilter option. In order to draw Figure 1 it was necessary to calculate 39 metric regression coefficients for successive achievement subgroups.

The syntax file with macros in it, enabled the performance of this task to be carried out efficiently. Without providing great detail, it would be of value to describe briefly the general construction of the syntax file that was used to generate the regressions coefficients. This syntax file, throughout a series of loops, allowed for the selection from the PISA data file cases for the required countries and for the required percentage groups and finally for the required variables. Calculated metric regression coefficients together with their standard errors and significance tests were merged and sent to Excel files. At each stage of developing the syntax file, the cross tests were undertaken to ensure that the obtained coefficients were correct. So, for example, Figure 2 shows the output from SPSS when a regression coefficient was calculated as a cross-test without using macros. The unstandardised regression coefficient ($b = 4.29$) from Figure 2 is equal to the value of the relevant point (T50) in Figure 1.

Table 1 Part of the file A1.1 with data, which were used to generate Figure 1 and with standard errors and significance tests (T-values are also reported) associated with the regression of Science achievement on Sex of Students

	b	SE	T	Sig.
Top 5	18.33	7.19	2.55	0.01
Top 10	12.81	5.38	2.38	0.02
Top 15	7.27	4.51	1.61	0.11
Top 20	7.61	3.94	1.93	0.05
Top 25	6.86	3.56	1.93	0.05
Top 30	7.57	3.3	2.29	0.02
Top 35	6.95	3.12	2.23	0.03
Top 40	7.54	3.01	2.50	0.01
Top 45	6.6	2.94	2.24	0.02
Top 50	4.29	2.89	1.48	0.14
Top 55	6.38	2.86	2.23	0.03
Top 60	5.17	2.85	1.81	0.07
Top 65	5.72	2.86	2.00	0.05
Top 70	6.04	2.88	2.10	0.04
Top 75	4.31	2.92	1.48	0.14
Top 80	4.56	2.98	1.53	0.13
Top 85	5.64	3.06	1.84	0.07
Top 90	4.76	3.15	1.51	0.13
Top 95	1.75	3.29	0.53	0.59
100 percent	-5.3	3.69	-1.44	0.15

Initial $N = 2851$

	b	SE	T	Sig.
100 percent	-5.3	3.69	-1.44	0.15
Bottom 95	-9.55	3.43	-2.78	0.01
Bottom 90	-8.63	3.36	-2.57	0.01
Bottom 85	-10.39	3.33	-3.12	0
Bottom 80	-11.04	3.33	-3.32	0
Bottom 75	-12.61	3.32	-3.80	0
Bottom 70	-12.1	3.34	-3.62	0
Bottom 65	-12.46	3.35	-3.72	0
Bottom 60	-11.57	3.39	-3.41	0
Bottom 55	-13.03	3.44	-3.79	0
Bottom 50	-14.3	3.51	-4.07	0
Bottom 45	-13.67	3.59	-3.81	0
Bottom 40	-15.96	3.69	-4.33	0
Bottom 35	-16.9	3.85	-4.39	0
Bottom 30	-15.19	4.06	-3.74	0
Bottom 25	-19.19	4.32	-4.44	0
Bottom 20	-21	4.79	-4.38	0
Bottom 15	-19.04	5.6	-3.40	0
Bottom 10	-17.79	6.96	-2.56	0.01
Bottom 5	-34.25	9.26	-3.70	0

Boys scored 2, Girls scored 1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	593.21	4.60		128.92	.000
	Sex - Q3	4.29	2.89	.039	1.48	.139

a. Dependent Variable: Warm estimate in Science (WLE)

Figure 2 Part of the output from SPSS generated without using the macro for the top 50 per cent of Australia's sample and variable Sex of Student

Another graph is shown in Figure 3 where an unstandardised regression coefficient is plotted for different achievement subgroups. Again there is an interesting relationship for the high-achieving students. In this case the variable, that is analysed, is 'Sense of belonging' and in order to provide a better understanding of it, the quotation from the *PISA 2000 Technical Manual* is presented[2].

In this example, as shown in Table 2, negative and significant regression coefficients are shown so, for example, in the top ten per cent of students those who did not have a strong sense of belonging to the school performed better than those students with a higher sense of belonging.

Table 2 Tests of significance for variable Sense of belonging and higher achieving subgroups of students in Figure 3

Top 5%	Top 10%	Top 15%	Top 20%	Top 25%	Top 30%	Top 35%	Top 40%	Top 45%	Top 50%	Top 55%
0.04	0.10	0.05	0.10	0.02	0.06	0.12	0.06	0.08	0.24	0.16

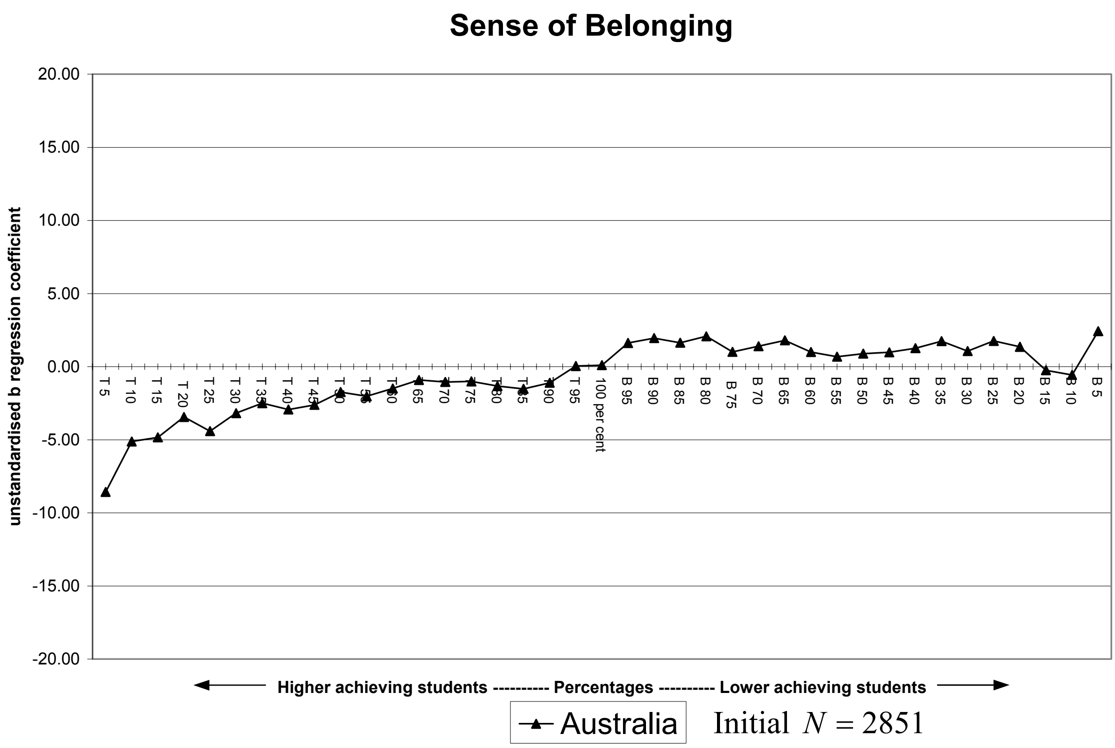


Figure 3 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable 'Sense of belonging' for different Science achievement subgroups from Australia (Source file A1.2). (NB Vertical scale is set from –20 to 20 because of later comparison between different countries)

The advantage of the graphs like those in Figure 1 or Figure 3, which can be called 'Percentage Population Plots' (PP-plots), is that successive unstandardised regression coefficients are presented to reveal a pattern associated with the successive levels of achievement of the subgroups of science students ranging from a small group (top 5 %) on the left-hand side of the

2 The PISA index of sense of belonging was derived from students' reports on whether their school is a place where they: feel like an outsider, make friends easily, feel like they belong, feel awkward and out of place, other students seem to like them, or feel lonely. A fourpoint scale was used with response categories: strongly disagree, disagree, agree and strongly agree. Scale scores are standardised Warm estimates where positive values indicate more positive attitudes towards school and negative values indicate less positive attitudes towards school.

(Adams and Wu, 2002, p. 229)

graph to the small group (bottom 5 %) on the right-hand side of the graph. While the extreme groups in the graphs are relatively small (about 140 cases in the case of the Australian sample) and may have sizable errors associated with the metric regression coefficients, the regression coefficients in the middle section of the graph have much smaller errors, since they are associated with large student groups. The statistical significance of the unstandardised regression coefficients, while estimated in Table 1 under the assumption of a simple random sample, does not take into consideration the changing design effects for each estimate, that necessarily increase the standard errors and increase p values associated with the tests of significance. However, the pattern of the graphs drawn in Figures 1 or 3 is highly informative, although the estimates of the unstandardised regression coefficients in the tails of graphs sometimes clearly indicate instability in the estimation procedure involved in the PP-plots.

The two above examples of graphs are presented merely to introduce the idea that there are differences between successive achievement groups in their regression relationships for particular variables. Because the main aim of this article is to introduce the PP-plots and show some possibilities of using these graphs, the issue of explaining why patterns of particular shapes occur is not developed any further at this stage.

Question 2: Does the same relationship hold across different student performance levels when comparing countries?

It is a well-known fact that there are differences between countries in the extent to which some factors influence student achievement; unfortunately most of the published findings report relationships for a whole population. It is mentioned above that the information available for Iran and Korea seems to show that both countries place great importance on supporting their more able students to take part in the International Physics Olympiads. On the contrary, there is a marked difference in the mean level of performance of the students between these two countries. Similarly there can be situations in which, for other pairs of countries, the students on average perform at the same level, but there are marked differences in the performance of the lower achieving students. Therefore, it is interesting to compare how different variables relate to Science achievement across different countries and across different achievement subgroups.

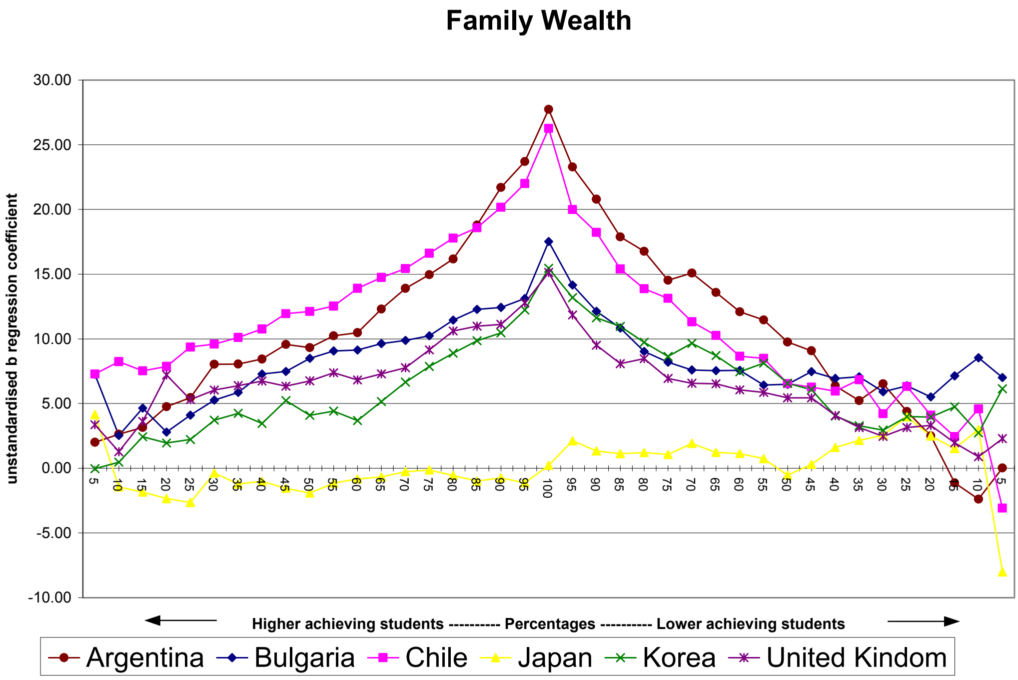
Family wealth

In Figure 4 the PP-plot is presented for six countries in which a PISA variable Family wealth[3] is argued to have a positive influence on student Science achievement, although with different values for the starting values of population estimates for the different countries. For five of the countries the path is symmetrically declining when moving towards higher and lower achieving students groups, except for Japan for which for almost all achieving subgroups Family wealth is not related to the Science achievement scores. For all countries shown in Figure 4, the PP-plots are roughly symmetrical when the left half of the graph is compared with the right half. Interestingly, this symmetrical relationship is not always shown for all countries. In Figure 5 a group of countries are presented, for which regression coefficients calculated between Family wealth and Science achievement scores are higher when moving towards better achieving students than when moving towards lower achieving subgroups of students. Moreover, in Figure 6 the opposite situation is shown.

---

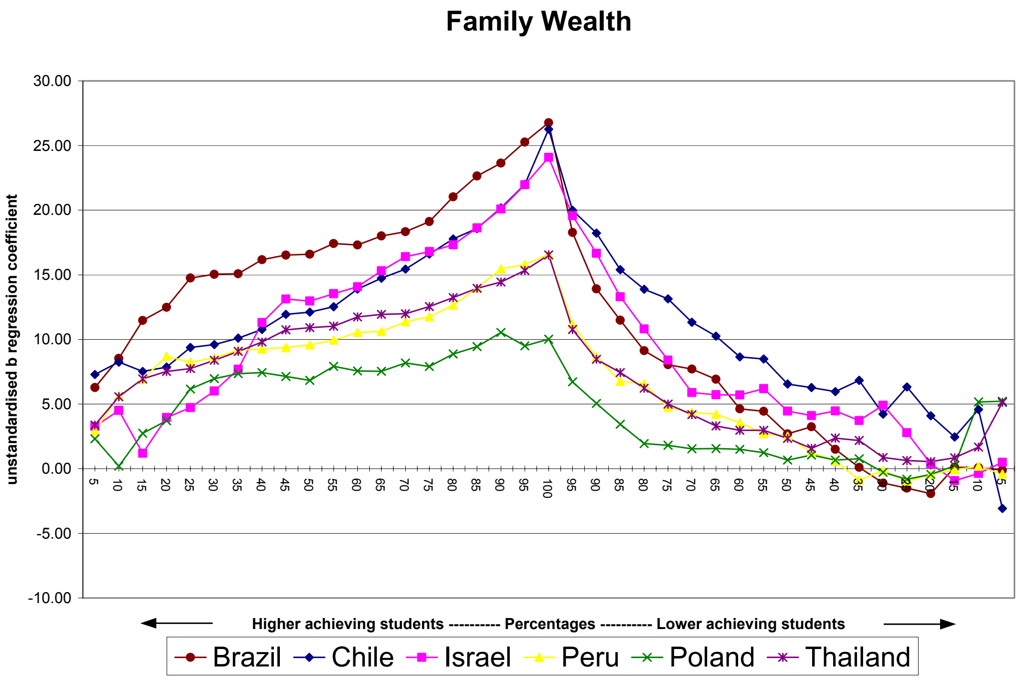
[3] The PISA index of family wealth was derived from students' reports on: (i) the availability in their home of a dishwasher, a room of their own, educational software, and a link to the Internet; and (ii) the numbers of cellular phones, televisions, computers, motor cars and bathrooms at home. Scale scores are standardised Warm estimates, where positive values indicate more wealth-related possessions and negative values indicate fewer wealth-related possessions.

(Adams and Wu, 2002, p. 224)



Initial	Initial	Initial	Initial	Initial	Initial
N =2206	*N* =2542	*N* =2703	*N* =2908	*N* =2757	*N* =5155

Figure 4 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable Family wealth for different Science achievement subgroups from six countries: Argentina, Bulgaria, Chile, Japan, Korea, and United Kingdom (Source file A1.3).



Initial	Initial	Initial	Initial	Initial	Initial
N =2701	*N* =2703	*N* =2397	*N* =2457	*N* =2043	*N* =2965

Figure 5 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable Family wealth for Science achievement subgroups from six developing countries: Brazil, Chile, Israel, Peru, Poland, and Thailand (Source file A1.3).

Clearly, when comparing Figure 5 and Figure 6, countries seem to be grouped with respect to their level of development. In the case of the more developed countries in Figure 6 there is a change in the regression coefficient from positive to zero or even to negative with movement towards the high achieving subgroups of students. On the one hand it can be argued that such a pattern occurs because in developed countries there is free education with small differentiation in teaching quality between schools, so that the students who want to study Science, have plenty of opportunities to do so, regardless of their family wealth. Moreover, the negative sign of the regression coefficients, which indicate that students from richer families obtain lower scores compared to students from poorer families, may be because richer students' parents do not encourage their children to study science, preparing them to study law, economics and commerce. Alternatively, it may be likely that a career in science related fields provides greater possibilities for upward social mobility that is sought by students from poorer families.

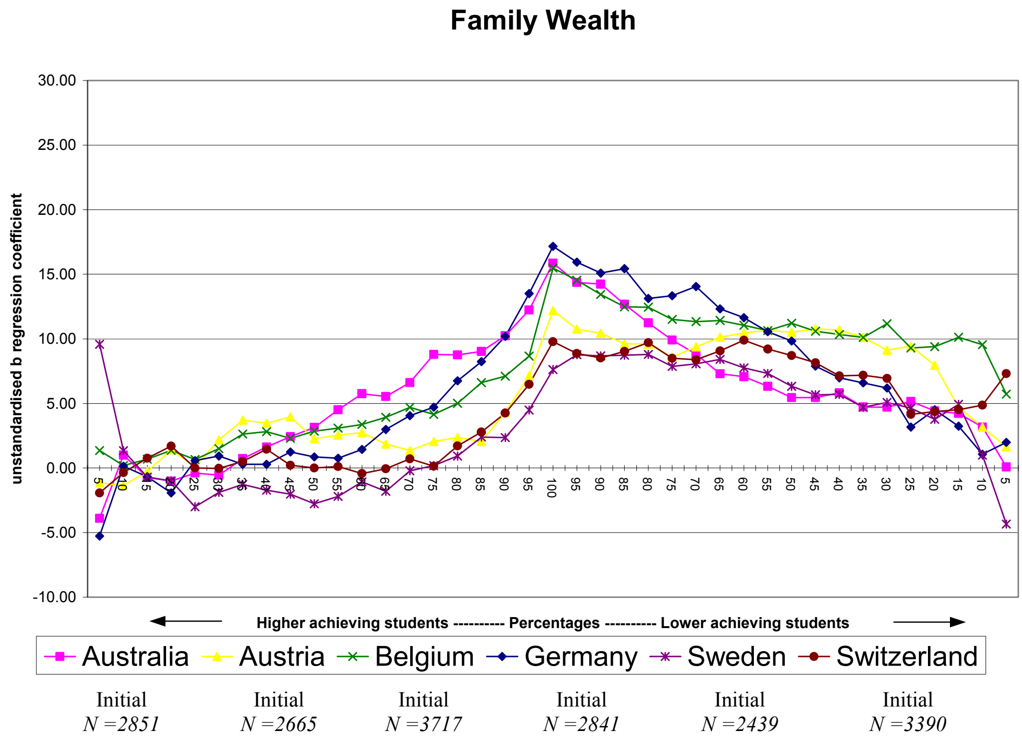


Figure 6 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable Family wealth for different Science achievement subgroups from six developed countries: Australia, Austria, Belgium, Germany, Sweden, and Switzerland (Source file A1.3).

On the other hand in Figure 5 an opposite relationship is shown for developing countries in which Family wealth is positively related to Science achievement for higher achieving students. This seems to indicate that there are greater career rewards in scientifically based occupations (e.g. Medicine) that are very attractive to students from wealthy homes in developing countries.

Scatter plots and the meaning behind PP-plots

It is useful to add an additional explanation that may help to provide a better understanding of the meaning behind PP-plots. In a sense a PP-plot provides more detailed information about the shape of the scatter plot that shows the relationships between Science achievement and, in the case considered above, Family wealth. In Figures 5a and 6a the scatter plots for two countries Brazil and Germany are presented in relation to their PP-plots in Figures 5 and 6 respectively for the complete sample. It can be seen in the case of Brazil that the regression line does not change substantially when compared to the complete sample, when the low achieving students are dropped. This shows that family wealth relates positively to the students' achievement even for the better students. However, the regression line becomes flatter if the high achieving students are deleted. That is clearly seen in the PP-plot in Figure 5 as well. When looking at the scatter plot for Germany (Figure 6a), it is seen to correspond to the PP-plot pattern. For example, for the high achieving students it is seen, that the regression line is flatter.

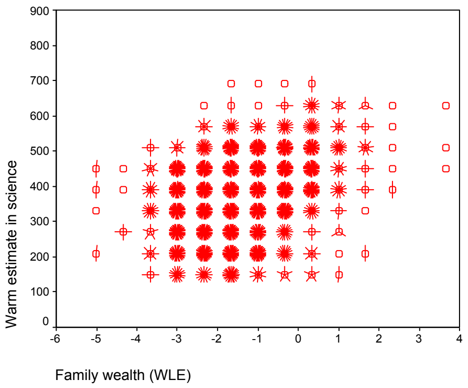


Figure 5a Scatter plot with relationship between Family wealth and Science achievement for Brazil

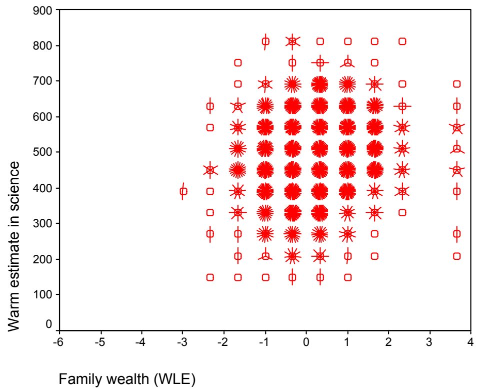


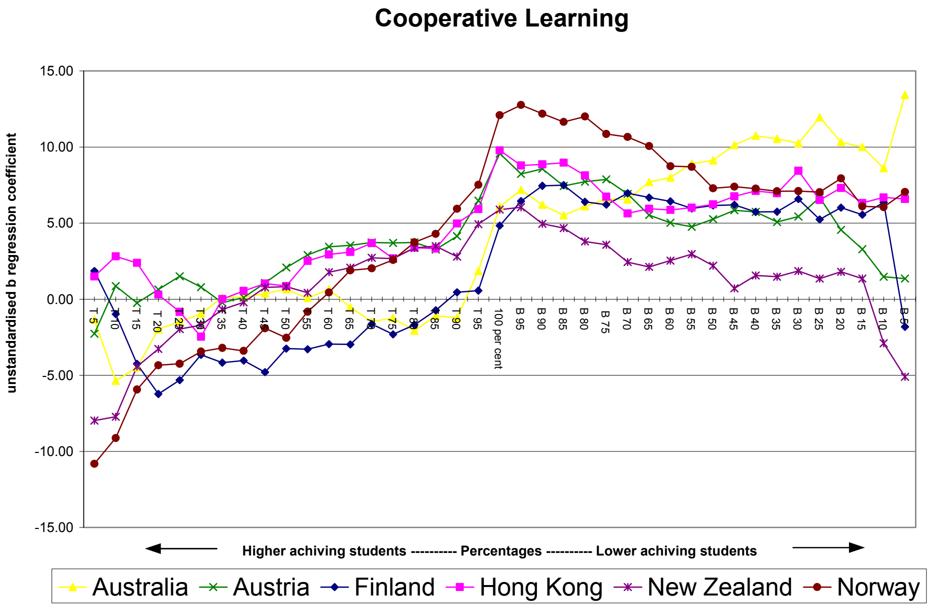
Figure 6a Scatter plot with relationship between Family wealth and Science achievement for Germany

Quite often in a research situation it is difficult to explain the relationships lying behind one particular regression coefficient in an estimated path model, even when many well established statistical methods are available. In this article a more general approach is presented for comparing regression coefficients for many countries and achievement subgroups. Perhaps it makes the task of examining the estimated relationships even more difficult. Many questions have to be asked even before starting, for example: How well does the Family wealth PISA variable reflect wealth in very poor countries? Therefore, explanations advanced are often highly speculative. Nevertheless, the main purpose of this article is to introduce the PP-plots and to examine such relationships further in order to understand better the meaning of the graphs. However, Figures 4, 5, 6 seem to address interesting so-called 'big picture' issues. The exception in the case of Japan shown in Figure 4 is of considerable interest.

Co-operative learning

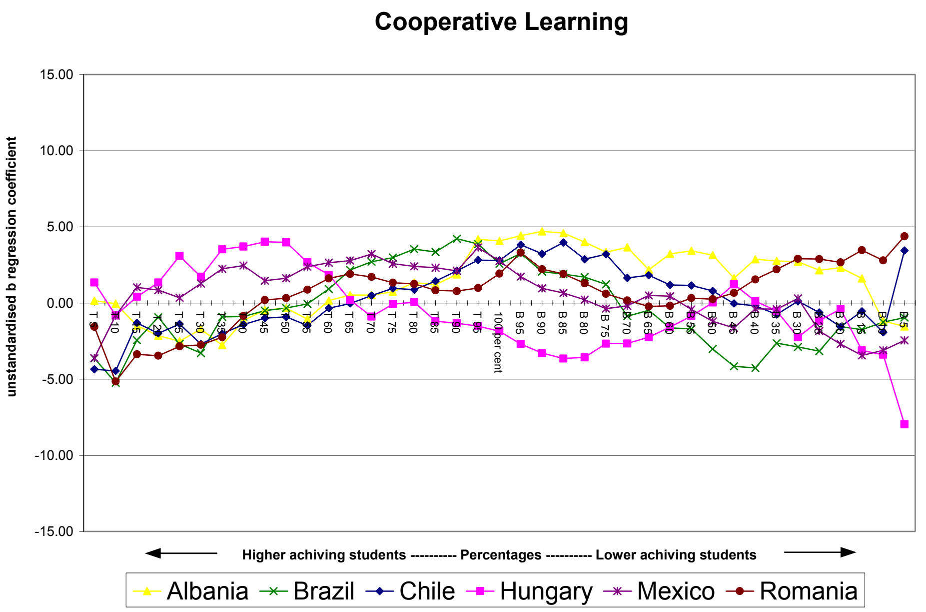
Another example of an interesting between countries grouping is shown when analysing the PP-plots for the variable, Co-operative learning[4]. PP-plots for two groups of countries are presented in Figures 7 and 8. For the countries in Figure 8 values of the regression coefficients are restricted to the range –5 and 5, showing that a self-perceived view about Co-operative learning does not relate to Science achievement. This may be due to the very limited use of co-operative learning techniques in educational curricula within those countries shown on Figure 8. On the contrary, Figure 7 it can be seen that in subgroups with lower achieving students, those students, who have a preference for co-operative learning, are doing better in science.

[4] The PISA index of co-operative learning was derived from student reports on the four items in Figure 64. A four-point scale with the response categories disagree, disagree somewhat, agree somewhat and agree was used. For information on the conceptual underpinning of the index, see Owens and Barnes (1992). Scale scores are standardised Warm estimates where positive values indicate higher levels of self-perception of preference for co-operative learning and negative values lower levels of self-perception of this preference. How much do you disagree or agree with each of the following? I like to work with other students, I learn most when I work with other students, I like to help other people do well in a group, It is helpful to put together everyone's ideas when working on a project. (Adams and Wu, 2002, p. 237)



Initial *N* =2851 | Initial *N* =2665 | Initial *N* =2706 | Initial *N* =2432 | Initial *N* =2024 | Initial *N* =2298

Figure 7 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable Co-operative learning for different Science achievement subgroups from six developed countries: Australia, Austria, Finland, Hong Kong (China), New Zealand, and Norway (Source file A1.4).



Initial *N* =2772 | Initial *N* =2701 | Initial *N* =2703 | Initial *N* =2795 | Initial *N* =2527 | Initial *N* =2694

Figure 8 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable Co-operative learning for different Science achievement subgroups from six developing countries: Albania, Brazil, Chile, Hungary, Mexico, and Romania (Source file A1.4).

Question 3: Can PP-plots help in detecting which variable is more generally intra-student or individually based and which is more generally inter-student or culturally based for all student achievement subgroups or for particular student achievement subgroups?

On the one hand, when looking at the PP-plots in Figures 4, 5 and 6, there are differences with regard to the extent to which the variable Family wealth is positively related to the Science achievement scores at the whole population level. There are even greater differences in the values and signs of the regression coefficients when considering the different achievement subgroups. Consequently it is meaningful to conclude that the relationship between Family wealth and Science achievement may depend on the kind of culture the students come from. On the other hand, in Figure 9 the PP-plots for the variable Sex of Students from all countries in PISA 2000 are presented. It may be argued that, although there are differences between the PP-plots, a general pattern seems to hold except in the left and right tails of the PP-plots among the most able and least able students. Therefore, it may be said that the way this variable relates to Science achievement outcomes is less culturally based and more individually based, or alternatively there is very little variability in the gender based societal differences between the countries involved and as a consequence in expected achievement in Science.

Probably because all the international surveys of students' knowledge and Science achievement, that have been conducted so far, have collected data about students' sex, similar PP-plots could be generated and may give clues to support or disprove the above assumption.

It is likely to be very useful from the policy makers' point of view, to know which factors influence students' Science achievement and are not due to cultural impact.

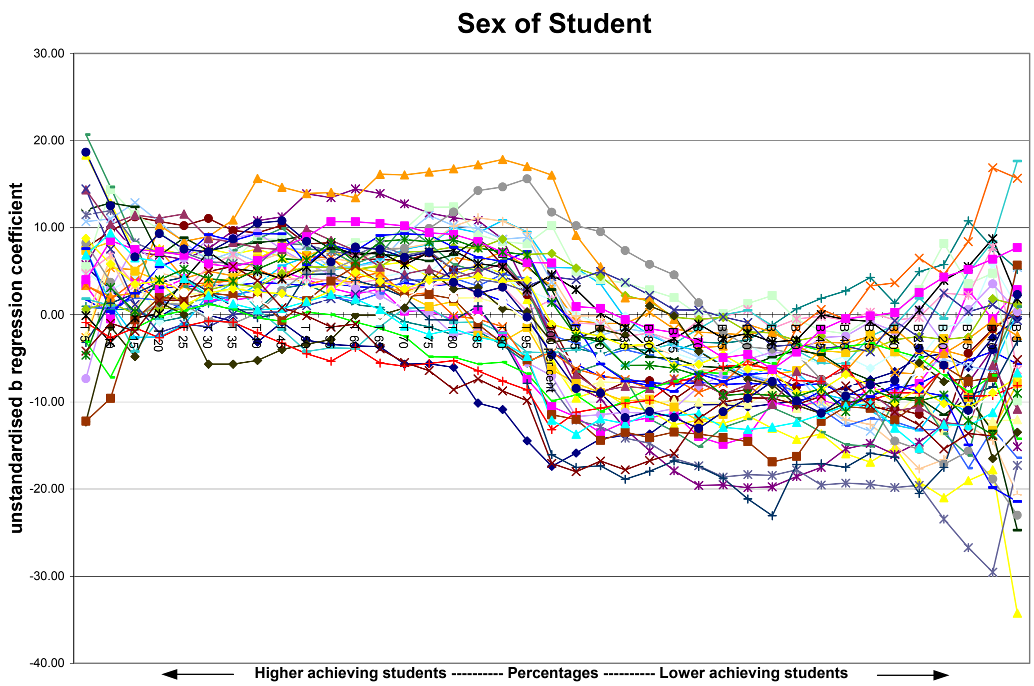


Figure 9 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable Sex of students for different Science achievement subgroups and for 42 countries (Source file A1.1).

Another general conclusion can be drawn from Figure 10. For 30 countries out of 33 it can be shown that the values of the regression coefficients of Science achievement regressed on

Academic self-concept[5] for the higher achieving subgroups of students are greater than those on the right-hand side of the PP-plots. This is not unexpected, because many recorded findings support the proposition that students with higher academic self-concept are achieving at a higher level than those with lower academic self-concept. Interestingly, the three countries which break the pattern are Brazil, Romania and Thailand, and the three with the highest PP-plots are Australia, Denmark and Sweden.

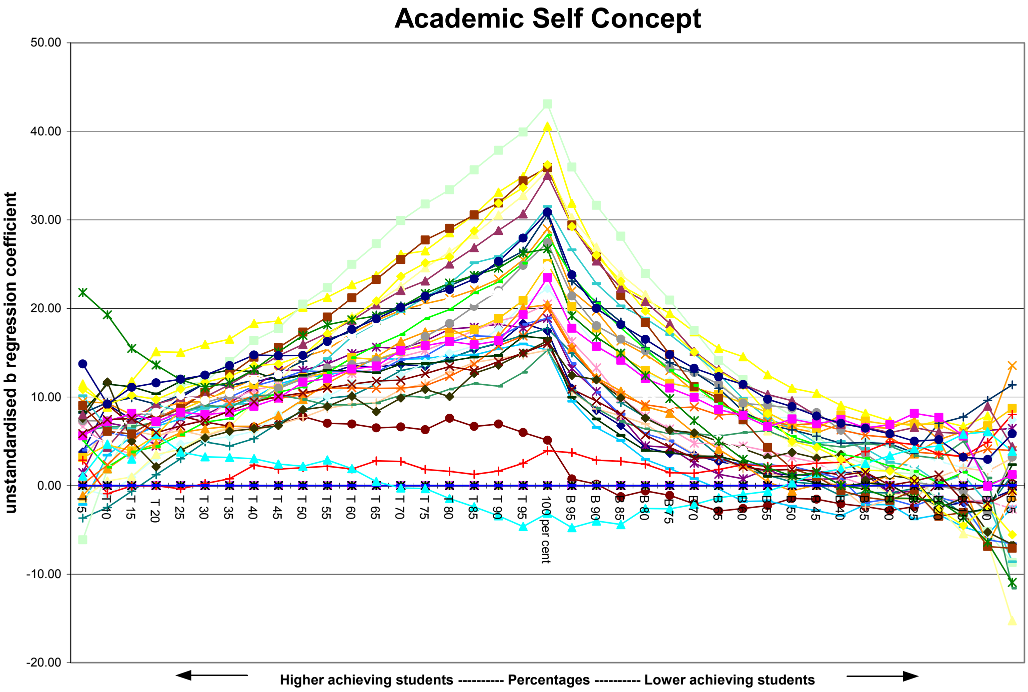


Figure 10 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable Academic self-concept for different Science achievement subgroups from 33 countries (Source file A1.5).

OTHER POSSIBILITIES ARISING FROM THE USE OF PP-PLOTS.

In a similar way, already developed syntax and macros with some small adjustments can be readily used with datasets from previous international surveys like IAEP, TIMSS, TIMSS-Repeat or PISA 2003. Obviously not all of the datasets collected are similar to the PISA 2000 set with respect to additional information from students, but for some variables, (for example, Sex of student) it is possible to generate PP-plots and to compare them with each other.

Moreover, generated in the way proposed, datasets with unstandardised regression coefficients can also be used in the meta-analyses. This possibility has not been developed further at this stage, but the regression coefficients from PISA2000 for whole national samples were used as a

---

[5] The PISA index of academic self-concept was derived from student responses to the items in Figure 68, which gives item parameters used for the weighted likelihood estimation. A four-point scale with the response categories disagree, disagree somewhat, agree somewhat and agree was used. For information on the conceptual underpinning of the index, see Marsh, Shavelson and Byrne (1992). Scale scores are standardised Warm estimates where positive values indicate higher levels of academic self-concept and negative values, lower levels of academic self-concept.

How much do you disagree or agree with each of the following?

I learn things quickly in most school subjects.

I'm good at most school subjects.

I do well in tests in most school subjects.

(Adams and Wu, 2002, p. 238)

base dataset for a Hierarchical Cluster Analysis. Interesting and perhaps to be expected, clustering of the countries on the basis of 20 variables (listed in Table 3) were available for all 42 countries and is presented in Figure 11. Three missing coefficients in the total dataset of the 837 coefficients were replaced with mean values. From Figure 11 it can be concluded that those 20 factors influence the Science achievement in a similar way for countries that are clustered closely together. A particularly strong and separate cluster is formed by Bulgaria, Czech Republic, Hungary and Poland. This cluster is an example, that may not have been predicted in advance, but when observed is highly meaningful and of considerable interest.

Table 3 The list of variables used in Hierarchical Cluster Analysis together with descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Sex	42	-17.40	16.03	-2.28	7.62
Mother international social and economical index	42	.50	1.95	1.30	.35
Father international social and economical index	42	.41	2.04	1.31	.41
Student self-expected international social and economical index	42	.24	2.36	1.51	.54
In. Socio-Econ. Index of father or mother	42	.34	2.07	1.31	.40
Father ISCED qualification	42	6.61	36.67	15.88	6.61
Mother ISCED qualification	42	5.63	35.01	16.93	6.39
Parental Academic interest (WLE)	42	4.25	26.53	15.64	5.58
Patental Social interest (WLE)	42	.81	19.07	8.75	4.43
Family educational support (WLE)	42	-17.49	8.78	-8.44	5.07
Family wealth (WLE)	42	-6.33	27.73	13.78	7.79
Home educational resources (WLE)	42	7.13	33.61	19.07	5.96
Cultural activities of students (WLE)	42	-3.84	30.39	14.75	7.95
Cultural possession of the family (WLE)	42	-.50	29.75	19.64	6.36
Time spent on homework (WLE)	42	-4.80	26.74	12.60	8.41
Teacher support (WLE)	42	-12.01	13.01	.52	5.66
School disciplinary climate (WLE)	42	-17.19	11.33	-6.05	5.83
Teacher-student relationship (WLE)	42	-8.44	14.91	3.43	7.77
Achievement press (WLE)	42	-14.05	11.33	-2.59	6.76
Sense of belonging (WLE)	42	-1.86	19.61	7.39	6.10

There is another possibility, although also not as yet developed, that may involve using the PP-plots to provide a way to group and classify countries. For example, the areas under the PP-plot curves for separate halves can be calculated and divided by each other. In this way an index for each country can be generated. In the case of the variable Family wealth such an index may provide information about how egalitarian particular countries are with respect to Science education. In the case of Academic self-concept such an index may help in the investigation of the degree to which academic self-confidence promotes higher achieving students compared to lower achieving in the learning of Science. There may be another advantage in the development of such an index. The same PP-plots for the same variable can be generated from the data collected for different international studies and allow for meaningful comparisons of the data collected in these studies.

```
* * * * * * H I E R A R C H I C A L   C L U S T E R   A N A L Y S I S * * * * * *

 Dendrogram using Average Linkage (Between Groups)
```

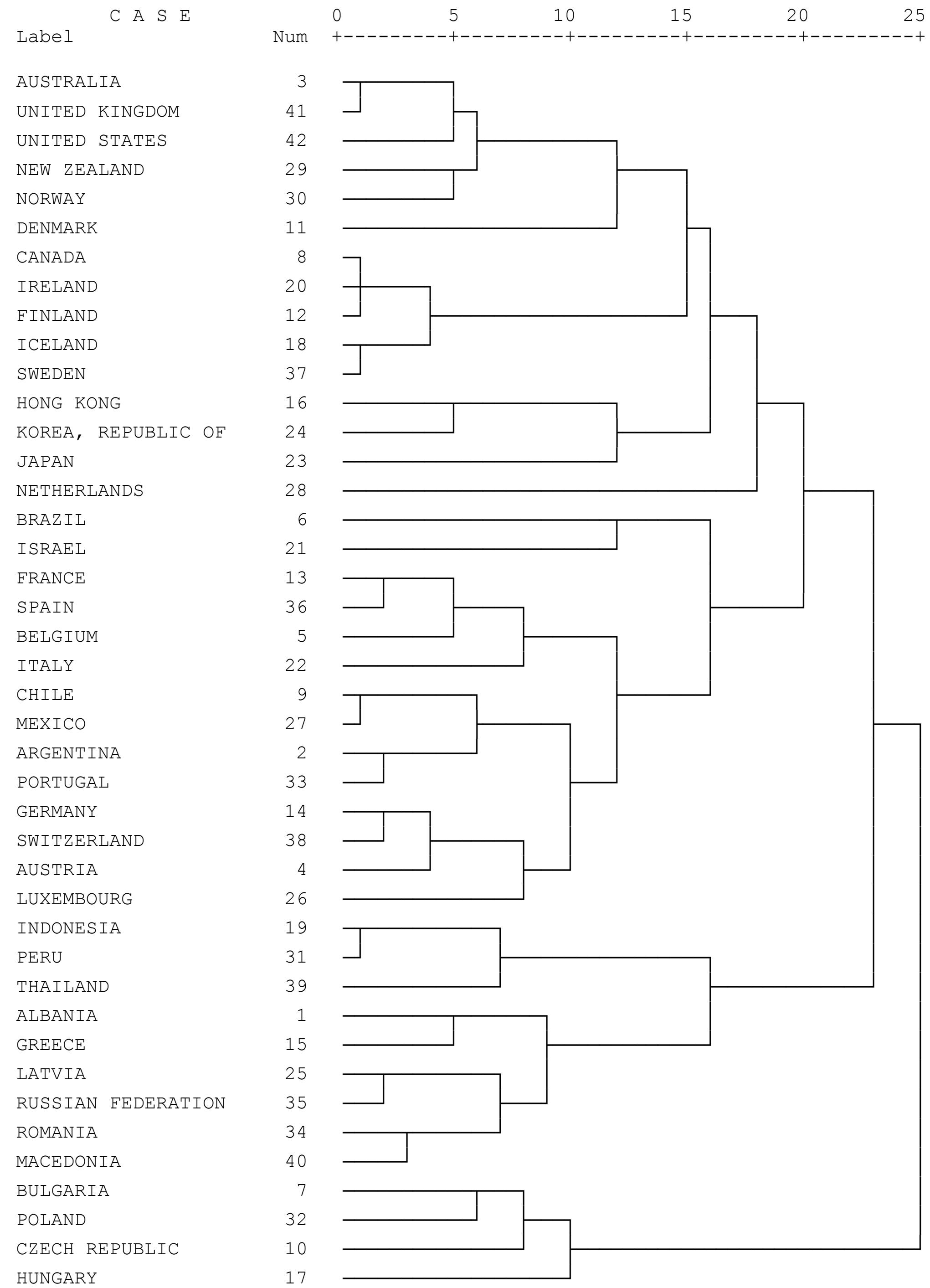
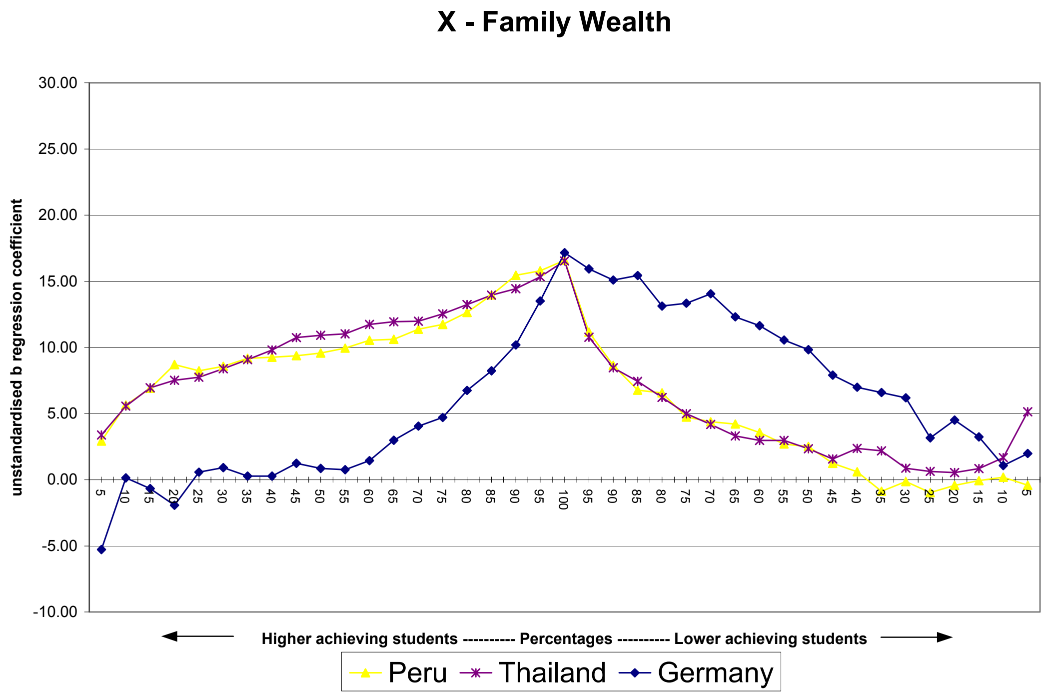


Figure 11 Dendogram generated after Hierarchical Cluster Analysis with unstandardised regressions coefficients for all national samples (File 1.6)

CONCLUSIONS

One question is of great interest. Do the research findings from one country apply to another country? This question is particularly important in the light of limited human and financial resources for educational research.

For example, for three countries from the PISA survey: A, B and C, Science achievement when regressed on variable X may yield similar values of a metric regression coefficient at the whole population level. It would not be enough though, to apply the policy conclusions from research in a field connected with the variable X, that were made in country C, to both countries A and B. However, when examining Figure 12, which present the PP-plot for the whole range of achievement levels, the graphs for countries A and C are very close to each other and both very different from the graph for country B. Would it be more justified to argue, on the base of similar PP-plot shapes, that a particular variable influences Science achievement in a similar way for these two countries? Would it be more legitimate to apply policy conclusions from research in a field connected with the variable X in an exchangeable way between countries A and C? This is a very simplified example, but these unanswered questions are of considerable importance in comparative research in the field of education, especially since Peru and Thailand are both developing countries with limited resources for research in education. However, because of the extensive body of educational research in certain highly developed countries, it is commonly assumed that similar relationships apply in developing countries.

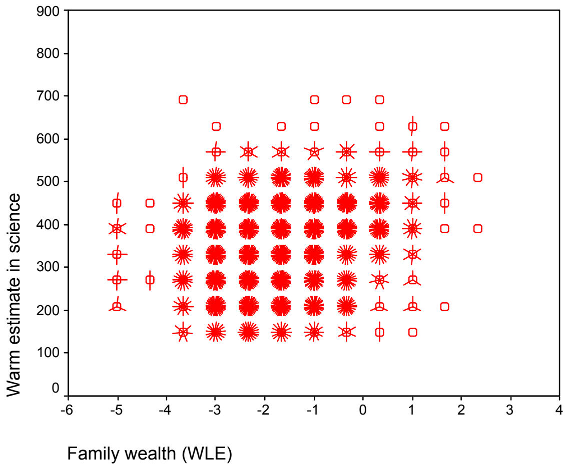


Initial *N =2457*　　Initial *N =2965*　　Initial *N =2841*

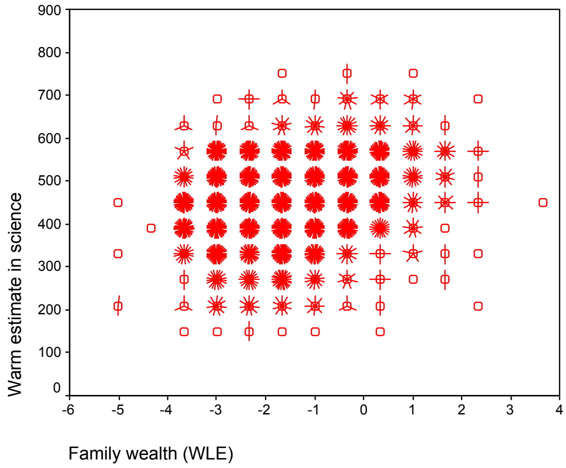
Figure 12 Plot of the unstandardised regression coefficients for Science achievement regressed on the independent variable X – Family wealth for different Science achievement subgroups and A – Peru B - Germany C - Thailand (Source file A1.3)

In the same way as is discussed in the Family wealth section, in addition to the PP-plots in Figure 12 the appropriate scatter plots were generated and are presented in Figure 13. The first two with very similar shapes were generated for Peru and Thailand and the third for Germany. The similarities between scatter plots for Peru and Thailand seem to be obvious and support those observed with PP-plots. However, the decision about the similarity between scatter plots is based on a subjective judgment, when PP-plots permit approaching this problem in a way that leads more readily to calculation.

PERU



THAILAND



GERMANY

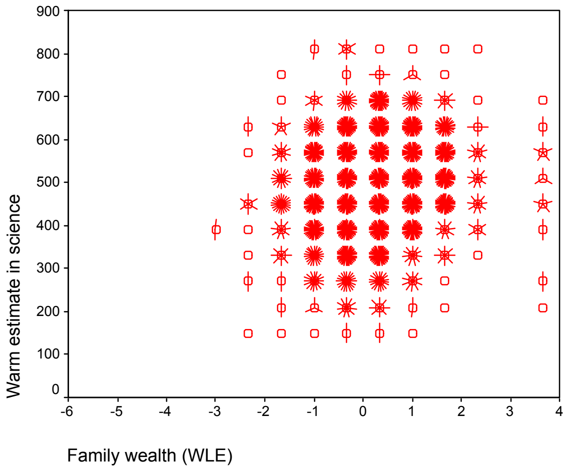


Figure 13 Scatter plots with relationship between Family wealth and Science achievement for Peru, Thailand and Germany respectively.

The main ideas that are at the forefront of a proposed strategy which involves the use of PP-plots can be stated in questions: How can countries be compared in the magnitude to which a particular variable influences Science achievement and how can this comparison be made across all achievement levels and the important subgroups of high and low performing students? In this article a new strategy has been introduced that may be the first step towards obtaining such a two dimensional comparison. It has to be noted, however, that through the PP-plots it is possible to investigate only how one variable at a time influences Science achievement, although it does examine data collected from different countries and for different achievement subgroups. This means that the many analytic possibilities that are available through using multivariate and multilevel analyses are not used here at all.

An interesting and important extension of the idea underlying the formation of fractiles using PP-plots, is to extend this idea to the analyses of simple multivariate and multilevel models that are tested initially with partial least squares programs which are robust under the conditions of lack of normality in the score distributions. It is highly probable that very different factors operate to influence both the achievement and attitudinal scores in science and mathematics of very high and very low performing students. These issues must be addressed in the cross-national testing programs in addition to the simplistic, although highly accurate estimation and ranking of national mean scores.

REFERENCES

Adams, R., & Wu M. (eds.) (2002). PISA 2000 Technical Report. Paris: OECD.

Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., Gregory, K. D., Smith, T. A., Chrostowski, S. J., et al. (2000). TIMSS 1999 International Science Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade. Chestnut Hill, MA: Boston College.

28th International Physics Olympiad. Sudbury, Ontario, Canada. 1997. [Online]. Available: http://laurentian.ca/physics/OLYMPIAD/PROCEED/RESULTS/INDEX.HTML#gold [2006 , April 11].

XXIX International Physics Olympiad. Taeknigardur, University of Iceland, Reykjavik. 1998. [Online]. Available: http://www.hi.is/pub/ipho/index.html [2006 , April 11].

XXX International Physics Olympiad. Padova, Italy. 1999. [Online]. Available: http://www.pd.infn.it/Olifis/welcome.htm [2006 , April 11].

XXXI International Physics Olympiad. University of Leicester, Leicester. 2000. [Online]. Available: http://www.star.le.ac.uk/IPhO-2000/results_silver.html [2006 , April 11].

XXXII International Physics Olympiad. Antalya, Turkey. 2001. [Online]. Available: http://www.ipho2001.org.tr/results/gold.html [2005 , January 22].

33rd International Physics Olympiad. Bali, Indonesia. 2002. [Online]. Available: http://www.fi.itb.ac.id/~ipho33/results.html [2006 , April 11].